Tran-SET
Transportation Consortium of South-Central States
Solving Emerging Transportation Resiliency, Sustainability, and Economic Challenges through the Use of Innovative Materials and Construction Methods: From Research to Implementation

Combining Virtual Reality and Machine Learning for Enhancing the Resiliency of Transportation Infrastructure in Extreme Events

Project No. 18ITSLSU09

Lead University: Louisiana State University

Final Report

September 2019

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Acknowledgements

We thank Christopher Melson, Dr. Husam Sadek, and Dr. Marwa Hassan for guidance throughout the project.

TECHNICAL DOCUMENTATION PAGE

1. Project No.	**2. Government Accession No.**	**3. Recipient's Catalog No.**
18ITSLSU09		

4. Title and Subtitle Combining Virtual Reality and Machine Learning for Enhancing the Resiliency of Transportation Infrastructure in Extreme Events	**5. Report Date** Sept. 2019 **6. Performing Organization Code**
7. Author(s) PI: Supratik Mukhopadhyay https://orcid.org/0000-0003-0839-1133 Co-PI: Yimin Zhu https://orcid.org/0000-0002-3871-4733 Co-PI: Ravindra Gudishala https://orcid.org/0000-0001-9767-1131	**8. Performing Organization Report No.**
9. Performing Organization Name and Address Transportation Consortium of South-Central States (Tran-SET) University Transportation Center for Region 6 3319 Patrick F. Taylor Hall, Louisiana State University, Baton Rouge, LA 70803	**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** 69A3551747106
12. Sponsoring Agency Name and Address United States of America Department of Transportation Research and Innovative Technology Administration	**13. Type of Report and Period Covered** Final Research Report Mar. 2018 – Mar. 2019 **14. Sponsoring Agency Code**

15. Supplementary Notes

Report uploaded and accessible at Tran-SET's website (http://transet.lsu.edu/).

16. **Abstract**

Traffic management models that include route choice form the basis of traffic management systems. High-fidelity models that are based on rapidly evolving contextual conditions can have significant impact on smart and energy efficient transportation. Existing traffic/route choice models are generic and are calibrated on static contextual conditions. These models do not consider dynamic contextual conditions such as the location, failure of certain portions of the road network, the social network structure of population inhabiting the region, route choices made by other drivers, extreme conditions, etc. As a result, the model's predictions are made at an aggregate level and for a fixed set of contextual factors. There is a clear need to develop traffic models that take into account local contexts and are closer to ground reality to provide government agencies the ability to make well-informed model-based decisions/policies.

In this project: (1) used Immersive Virtual Environment (IVE) tools for generating context-aware and high-fidelity data related to drivers' route choice behavior, (2) developed a novel approach for developing high-fidelity route choice models with increased predictive power by augmenting existing aggregate level baseline models with information on drivers' responses to contextual factors obtained from stated choice experiments carried out in an IVE through the use of knowledge distillation. To this end, the study used a virtual driving environment designed based on I-10 in Baton Rouge, LA. Five alternate routes were introduced to the participant. Ten experimental scenarios were conducted to produce initial data about drivers' dynamic route choice behavior, given emerging contextual factors. Experimental results have demonstrated that the predictions of the augmented models produced by our approach are much closer to reality than that of the baseline. Our study demonstrates that existing route choice models based on econometric theories cannot accurately predict behavior in real world scenarios. For high-fidelity route choice models, one needs to combine existing route choice models with information about contextual factors gleaned from SCEs.

17. Key Words	**18. Distribution Statement**
Interactive Virtual Environment, Machine Learning, Knowledge Distillation	No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	**20. Security Classif. (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified	Unclassified	20	

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized.**

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS, ABBREVIATIONS, AND SYMBOLS

IVE	Interactive Virtual Environment
MLP	Multilayer Perceptron
SCE	Stated Choice Experiment
VR	Virtual Reality

EXECUTIVE SUMMARY

Route Choice Models predict the route choices of travelers traversing an urban area. Most of the route choice models link route characteristics of alternative routes to those chosen by the drivers. The models play an important role in prediction of traffic levels on different routes and thus assist in development of efficient traffic management strategies that result in minimizing traffic delay and maximizing effective utilization of the transport system. High fidelity route choice models are required to predict traffic levels with higher accuracy. Existing route choice models do not take into account dynamic contextual conditions such as the occurrence of an accident, the socio-cultural and economic background of drivers, other human behaviors, the dynamic personal risk level, etc. As a result, they can only make predictions at an aggregate level and for a fixed set of contextual factors. For higher fidelity, it is highly desirable to use a model that captures significance of subjective or contextual factors in route choice.

The objectives of the project are:

- identify contextual factors that affect drivers' decisions;
- experiment the effectiveness of modeling such contextual factors in interactive virtual environment (IVE); and
- test the integration of machine learning with results from IVEs to improve predictions.

Results of this project will allow researchers in the future to further develop a novel context-aware framework that combines virtual reality with machine learning to predict:

- "optimal" routing of traffic under both normal and abnormal conditions (hurricanes, disasters, football games, etc.) minimizing the average driving time; and
- appropriate strategic allocation and placement of resources (scheduling traffic light, deploying personnel, sensors, sign boards, actuators, and materials).

The following provides a summary of the results:

Data Collection from IVE: The driving environment for this study is designed based on the I-10, starting off the Mississippi River bridge all the way to the College Dr. Along the way, five alternate routes were introduced to the participant in an IVE—Exits a, b, c, d, and e, the latter of which would be College Drive. Ten experimental scenarios were conducted to produce initial data about drivers' dynamic route choice behavior, given emerging contextual factors. These scenarios varied between normal, medium, and heavy traffic, combined with journey type and social impact. Every participant was exposed to all the driving scenarios including a baseline scenario. In the baseline scenario, we would collect information about participant's route choice pattern in normal traffic and non-urgent bound condition. Each scenario contained 1, 2, or 3 contextual factors. The first contextual factor was the traffic density which was varied over three levels, i.e. normal, medium, and heavy. The density thresholds were designed based on the Highway Capacity Manual (HCM) in freeways. The next factor was the purpose of the trip (journey type) which consisted of urgent-bound and non-urgent-bound trips; on the urgent-bound trip, participants were told to consider how important was it to meet the time of arrival commitment, while the non-urgent-bound posed no rush to reach the destination. The third factor in this study was the impact of other drivers' route selection; whether the participants' route choice would be influenced by observing other drivers taking an exit or not.

The study received approval from the IRB at Louisiana State University. Forty-one individuals (20 male and 21 females; age: 31.44 ± 7.97) volunteered to participate in the experiment. Prior to the experiment, participants were presented with a questionnaire asking the following items: (1) demographic characteristics (age, gender, race, education, employment status), (2) top concerns while they were stuck in the traffic congestion. Their choices included hours of extra travel time, speed reduction, monetized value of delay, additional vehicle operating cost, (3) familiarity with the area, and (4) socio-economic status (having concerns about spending less money on your gas). After answering the questionnaires, participants were asked to sit on a stationary chair at a desk with a driving wheel which was placed in front of a flat screen monitor where the driving simulation would run. Next, they were invited to practice for a few minutes to get acquainted with driving the simulator. After enough practicing with the driving simulator and becoming comfortable with its environment the research team would assign the participant to the scenarios (see below). The origin and the destination in all the scenarios were same and each scenario took about two minutes to finish. In each scenario different contextual factor(s) were presented and participants were required to choose their preferred route. Each participant was exposed to all the driving scenarios including a baseline scenario. Dynamic route guidance was presented to the participants where a driver is guided on to routes that will minimize travel time for the overall road network. Moreover, traffic noise was incorporated into the IVE setting and the audio was intended to help participants feel disconnected from the lab environment and be more present in the experiment. There were also textual cues to attract participants' attention to the spatial, temporal, and some other qualities of the scenarios.

The scenarios were counterbalanced and played out in a random fashion to avoid behavioral biases due to the order effect. The structure of the frequency of the route choice across each experimental scenario significantly varies and the p-value associated with the chi-square test confirms this observation ($p<0.001$, $\alpha=0.05$). The study, in particular, examines individuals' diversion tendency onto alternate routes that are induced by traffic condition, journey type, and the impact of social influence (seeing other drivers exiting, for example) while driving on the Interstate 10 (I-10) freeway in Baton Rouge, between the Mississippi River Bridge and College Drive exit.

Logistic Regression Model: In the case of our study, the structure of the frequency of the route choice across each experimental scenario significantly varies and the p-value associated with the chi-square test confirms this observation ($p<0.001$, $\alpha=0.05$). Thus, the route choice variable has an association with the contextual factors. Furthermore, we also sought to investigate the impact of each of the contextual factor as well as the human-related variables, by developing a predictive statistical model. To that end, a specific type of logistic regression, namely, logistic regression with GEE (Generalized Estimating Equation) was applied for this specific dataset for two main reasons: (1) the dependent variable of the study was dichotomous and a set of explanatory variables were available AND (2) the data were clustered, involved repeated measures. That is, 10 repetitions for each respondent were grouped as a cluster in a regression. From the model, we inferred that "traffic condition," "age," and "race" were statistically significant for determining route choice. In particular: (1) participants (drivers) tend to stay longer in I-10 if the traffic condition is normal or even medium, (2) older people tend to drive longer distances before exiting I-10, and (3) a driver with Middle Eastern background tend to exit I-10 more frequently rather than drivers with Asian background.

Machine Learning Model: We developed a novel approach for building high-fidelity route choice models with increased predictive power by augmenting existing aggregate level baseline models with information on drivers' responses to contextual factors obtained from stated choice experiment (SCE) carried out in an IVE through the use of knowledge distillation. Our approach uses the prior knowledge acquired by a teacher neural network pre-trained on data about drivers' responses to contextual factors to augment a student neural network (a baseline model) in a guided fashion. We demonstrated experimentally that the predictions of the augmented model are much closer to reality than that of the baseline.

1. INTRODUCTION

It is widely known that traffic congestion has significant environmental, economic, and public health consequences. The total cost and time loss associated with traffic congestion in the U.S. has been reported to be more than $121 billion per year and 38 hours per person, respectively. In the U.S., people mostly prefer to use freeways and highways, however in case of traffic congestions, alternative routes are also taken to avoid travel delay.

Mainstream research shows growing interest and need for better understanding drivers' route choice behavior *(1)* on model estimation and flow prediction. Route choice behavioral inputs are key in making accurate predictions about the patterns of traffic and providing real-time traffic information. According to Ben-Akiva et al. *(1)*, the choice set of alternative routes is one of the major components of the route choice models. In fact, route choice is a decision by drivers on a daily basis and the selection among a given choice set can be influenced by factors, such as road condition and human-related factors (driving experience, driver's socio-economic characteristics, and driving behavior and attitudes). Yet, collecting datasets that are sensitive to the aforementioned factors are challenging and the existing approaches usually take into account the general factors motivating drivers' route choice behavior. Sometimes, the data is collected at individual level but then aggregated to be able to develop a viable statistical model that reflects average behavior. The developed model is then applied at individual level and finally the results are aggregated to forecast future behavior. Most of the route choice models are currently using aggregated data. The current route choice models are calibrated using static contextual conditions and are not generally able to account for dynamic travel time, accessibility to the nearest freeway, traffic incidents, and road closures due to emergencies *(2)*. As a result, such models can only make predictions at an aggregate level and for a generic set of contextual factors. Apparently, the predictions that are made without considering such information would have considerable uncertainties. This study suggests the use of interactive virtual environment (IVE) tools to enable the incorporation of the contextual factors in data collection. The authors believe that high-fidelity models that are based on rapidly evolving contextual conditions can have a huge impact in the design and implementation of smart and energy-efficient transportation system. This study reports on experimental scenarios in which specific contextual factors are added in the testing design, using IVE platform and a driving simulator. The study, in particular, examines individuals' diversion tendency onto alternate routes that are induced by traffic condition, journey type, and the impact of social influence while driving in the Interstate 10 (I-10) freeway in Baton Rouge, between the Mississippi River Bridge and College Drive exit. Not only was the goal to elicit information about their route choice behavior, but also to learn about drivers learning behavior/or willingness to adapt to dynamic or emerging network conditions. The conditions surrounding each trip were varied to elicit preferences about values they place on factors that are believed to influence driving behavior. For instance, how willing drivers would be to switch their routes given that they are not expected to arrive at a destination at predetermined time versus when they are expected to meet a time commitment.

1.1. Problem Statement

Route choice models form the basis of traffic management systems. High Fidelity models that are based on rapidly evolving contextual conditions can have a huge impact on smart and energy efficient transportation. Existing route choice models are generic and are calibrated using static contextual conditions. The models do not take into account dynamic contextual conditions such as dynamic travel time, accessibility to nearest freeways, traffic incidents, and road closure due to an emergency. As a result, they can only make predictions at an aggregate level and for a generic set of contextual factors. There is a clear need to develop route choice models that take into account local contexts and are closer to ground reality to provide government agencies the ability to make well-informed model-based decisions and policies.

1.2. Motivating Scenario

In the event of a flooding or a storm (an extreme event) it is often the case that it results in the failure of certain links of a road network. This puts regular drivers of the area in a new route choice decision-making context when traveling between any given origin and destination. For example, a driver in Denham Springs heading to LSU campus may not be aware that a portion of I-12 is submerged in water and thus may not know what alternative routes he might consider ahead of time. But as the situation unfolds and the driver runs into the road closure situation, it puts the driver in a choice context where many factors are evolving in real time. For example, many dynamic and emergent contextual conditions such as remaining time for travel, familiarity with the area, personality traits (risk taking or risk averse), and the proximity to nearest alternative routes, gas stations with gasoline stocks, constitute a decision-making environment that is very different from when the driver starts the trip. A better understanding of factors influencing the driver's decision on spot and messages delivered to drivers for optimizing road network conditions is critical to managing traffic streams in the road network.

Routine route choice models that capture decision making rules of day-to-day commute do not reflect decision making that might occur in extreme events. In particular, they are generic and do not capture the contextual factors influencing driving behavior/decision-making in a given situation. Contextual factors and their impact on driver decision-making can be difficult to understand due to lack of data in the real world. Thus, it becomes imperative to use virtual reality environment to portray alternative scenarios by varying critical contextual factors and capture decision making rules that might occur in a new decision making context. Using IVE one can create possible decision making contexts that might arise in extreme scenario. The data on contextual factors collected from IVE experiments can be used to train machine learning engines to improve the predictive power of existing models for traffic routing and resource allocation and deployment of resources (sensors, personnel, etc.).

In order to make road networks resilient it is important to consider all possible potential link failures and thus routes and plan ahead for all contingencies. This might eventually help in making the road network resilient and less vulnerable in the event of an extreme storm or flooding.

2. OBJECTIVES

2.1. Purpose

The main purpose of this study is to establish a context-aware IVE experimental setting to generate first-hand data related to the individuals' route choice. The data that is gained through this approach will be sensitive to the design, its context and specific events. The study aims at incorporating the context-aware information gained from the IVE platform, aggregate it using statistical tools, and make relevant inferences about choice set given certain traffic conditions. This method of data collection also provides possibility of incorporating the influences of particular individual (human-related) variables into the route choice models and provides a more customizable research tool for evaluations and future predictions. Ultimately, this study provides a powerful computation and analytic framework that integrates machine learning-based models with Immersive Virtual Environments (IVEs) to improve the predictive power of existing models for traffic routing and resource allocation and deployment of resources.

2.2. Technical Objectives

Context-aware data-driven route choice models can enable efficient routing of traffic as well as strategic deployment of resources (personnel, materials, sensors, and actuators). It is important to understand how the use of such resources will affect driver's route selection decision, and support context-aware and driver-centered interactions to help drivers make proper decisions. Existing techniques for allocating and deploying resources (sensors, personnel, materials, etc.) are either based on econometric and game-theoretic approaches or are based on predictive models based on historical data. These models do not capture the contextual factors influencing driving behavior/decision-making in a given situation. In many cases, such approaches and models suffer from performance gaps: there is a significant gap between their predictions and the ground realities that is human-centered.

Existing predictive models do not take into account the contextual aspects in which drivers make decisions and thus influence the use and state of traffic infrastructures. As interactions between humans and traffic infrastructures are context driven, the lack of inclusion of specific contexts is a major source of performance gaps. Examples of contextual factors include the nature and characteristics of an extreme event, purpose of the trip, roadway conditions, communication infrastructure, characteristics of drivers, and social, climatic, and economic conditions.

The overarching goal of this project is to develop a powerful computation and analytic framework that integrates machine learning-based models with immersive virtual environment to improve the predictive power of existing models for traffic routing and resource allocation and deployment of resources (sensors, personnel, etc.) by taking into account contextual factors affecting human interaction with highway infrastructure. To achieve the goal, the project team will:

- identify contextual factors that affect drivers' decisions;
- experiment the effectiveness of modeling such contextual factors in IVE; and
- test the integration of machine learning with results from IVEs to improve predictions.

The first objective is to establish a context-aware IVE experimental setting to generate first-hand data related to the individuals' route choice. The data that is gained through this approach will be sensitive to the design, its context and specific events. The study aims at incorporating the context-aware information gained from the IVE platform, aggregate it using statistical tools, and make relevant inferences about choice set given certain traffic conditions.

Results of this project will allow researchers in the future to further develop a novel context-aware framework that combines virtual reality with machine learning to predict:

- “optimal” routing of traffic under both normal and abnormal conditions (hurricanes, disasters, football games, etc.) minimizing the average driving time,
- appropriate strategic allocation and placement of resources (scheduling traffic light, deploying personnel, sensors, sign boards, actuators, and materials).

The second goal of this study is to develop a novel approach for developing high-fidelity route choice models with increased predictive power by augmenting existing aggregate level baseline models with information on drivers' responses to contextual factors obtained from Stated Choice Experiments carried out in an Immersive Virtual Environment through the use of knowledge distillation.

3. LITERATURE REVIEW

It is widely known that traffic congestion has significant environmental, economic and public health consequences. The total cost and time loss associated with traffic congestion in the U.S. has been reported to be more than $121 billion per year and 38 hours per person, respectively *(3)*. In the U.S., people mostly prefer to use freeways and highways, however, in the case of traffic congestion, alternative routes are also taken to avoid travel delay *(4,5)*.

3.1. Route Choice Models

Mainstream research shows growing interest and needs for better understanding drivers' route choice behavior on model estimation and flow prediction *(1,6)*. Route choice behavioral inputs are key in making accurate predictions about the patterns of traffic and providing real-time traffic information. According to Ben-Akiva et al. *(7)* the choice set of alternative routes is one of the major components of the route choice models. In fact, route choice is a decision by drivers on a daily basis *(1)* and the selection among a given choice set can be influenced by factors, such as road condition and human-related factors— driving experience, driver's socio-economic characteristics and attitudes *(8)*. Yet, collecting datasets that are sensitive to the aforementioned factors are challenging and the existing approaches usually take into account the general factors motivating drivers' route choice behavior. Sometimes, the data is collected at an individual level but finally, the results are aggregated to develop a viable statistical model that reflects average future behavior. Currently, most of the route choice models are merely using the aggregated data.

The existing route choice models are calibrated using static contextual conditions and are not generally able to account for dynamic travel time, accessibility to the nearest freeway, traffic incidents, and road closures due to emergencies. As a result, such models cannot make predictions at fine levels and for a generic set of contextual factors. Apparently, the predictions that are made without considering such information would have considerable uncertainties. Transportation engineers have been studying commuter route choice behavior for four decades now. Engineers developing route choice models theorized that travel time plays a crucial and important role in the selection of a route. Route choice behavior theories began to evolve in the late eighties and early nineties as engineers' understanding of route choice behavior improved by studying data about empirical route choice behavior. Pursula and Talvite *(9)* developed a mathematical route model by postulating that drivers do consider other factors apart from travel time in making a route choice. In *(10, 11),* the authors discovered that commuters prefer to use habitual routes when traveling in familiar areas as opposed to choosing a route that provides them with maximum utility. Other researchers such as Doherty and Miller *(12)* investigating route choice found that apart from travel time, factors such as residential location, familiarity with the route, and employment locations are significant in the route choice process. Deep learning techniques *(13)* can be used to predict traffic congestion and route choice. However, deep learning models, being opaque, cannot be used to causally explain drivers' route choice.

In reviewing the existing research it can be gleaned that transportation researchers have employed two different types of empirical data collection in studying route choice behavior. First, collecting route choice data using observed actual choices and second, collecting route choice data in hypothetical experiments. Researchers have for the majority of cases used utility maximizing theory to explain route choice behavior that is rooted in econometrics *(7).*

Existing route choice models do not take into account dynamic contextual conditions such as the occurrence of an accident, the socio-cultural and economic background of drivers, other human behaviors, the dynamic personal risk level, etc. As a result, they can only make predictions at an aggregate level and for a fixed set of contextual factors.

3.2. IVE Platforms

IVE platforms have been useful for providing safe and user-friendly experimental settings, being inexpensive and highly portable *(14)*, as well as generating context-aware and high-fidelity data *(15)*. There have been several researches (e.g. *(16)* and *(17)*) using IVE to provide a realistic simulated setting for testing driving behaviors under various experimental conditions such as high-density traffic, fatigued, and drug-impaired. IVEs have also been applied to studies related to human-building interactions and energy usages. For instance, Heydarian et al *(18)* studied occupant lighting preferences in a single office using IVEs. Saeidi et al. *(19)* validated occupant light use behavior in IVEs and showed that IVEs were capable of replicating field experiences. Niu et al. *(20)* developed a framework to integrate building designs with IVEs to help building designers capture occupant preferences and identify context patterns. It is believed that high-fidelity models that are based on rapidly evolving contextual conditions can have a huge impact in the design and implementation of a smart and energy-efficient transportation system *(21*). IVEs have many limitations such as short experiment sessions, small data samples, and negative impacts on participants (e.g., cybersickness) *(22),* which make IVE-based experiments limited.

3.3. Machine Learning Techniques

Learning approaches based on SCE for estimating route choice have been attempted by *(23).* Deep learning techniques have achieved success in a variety of tasks *(24-29)*. Transportation engineers have used machine learning-based high resolution satellite imagery analysis *(24-26)* for mapping urban and rural regions in the process of designing satellite imagery. However, information gleaned from analyzing such imagery data only provides guidance at an aggregate level. To the best of our knowledge, the incorporation of context factors together with fine-grained route choice predictions at the individual level has not been done before.

4. METHODOLOGY

4.1. Purpose

The first goal of this study is to establish a context-aware VR experimental setting that can generate first-hand data related to the individuals' route choice. The data that is collected through this approach will be sensitive to the design, its context, and specific events. This method of data collection also enables to incorporate the influences of particular individual (human-related) variables into predictive models and provides a more customizable research tool for evaluations and future predictions. Eventually, this approach makes possible to develop a powerful computation and analytical framework that integrates machine learning-based models with immersive virtual environments (IVEs) which can greatly improve the predictive power of existing models for traffic routing, resource allocation, and deployment of resources.

4.2. Method

4.2.1. Design

The driving environment of this study is designed based on the I-10, starting off the Mississippi River bridge all the way to the College Dr. Along the way, five alternate routes (shown in Figure 1.) were introduced to the participant—Exits a, b, c, d, and e, the latter of which would be College Drive. Ten experimental scenarios were conducted to produce initial data about drivers' dynamic route choice behavior, given emerging contextual factors. These scenarios varied between normal, medium, and heavy traffic, combined with journey type and social impact (see Table 1). Every participant was exposed to all the driving scenarios including a baseline scenario. The baseline scenario would collect information about participant's route choice pattern in normal traffic and non-urgent bound condition. Each scenario contained 1, 2, or 3 contextual factors. The first contextual factor was the traffic density which was varied over three levels, i.e. normal, medium, and heavy. The density thresholds were designed based on Highway Capacity Manual (HCM) in freeways. The next factor was the purpose of the trip (journey type) which consisted of urgent-bound and non-urgent-bound trips; on the urgent-bound trip, participants were told to consider how important was it to meet the time of arrival commitment, while the non-urgent-bound posed no rush to reach the destination. The third factor in this study was the impact of other drivers' route selection; whether the participants' route choice would be influenced by observing other drivers taking an exit or not.

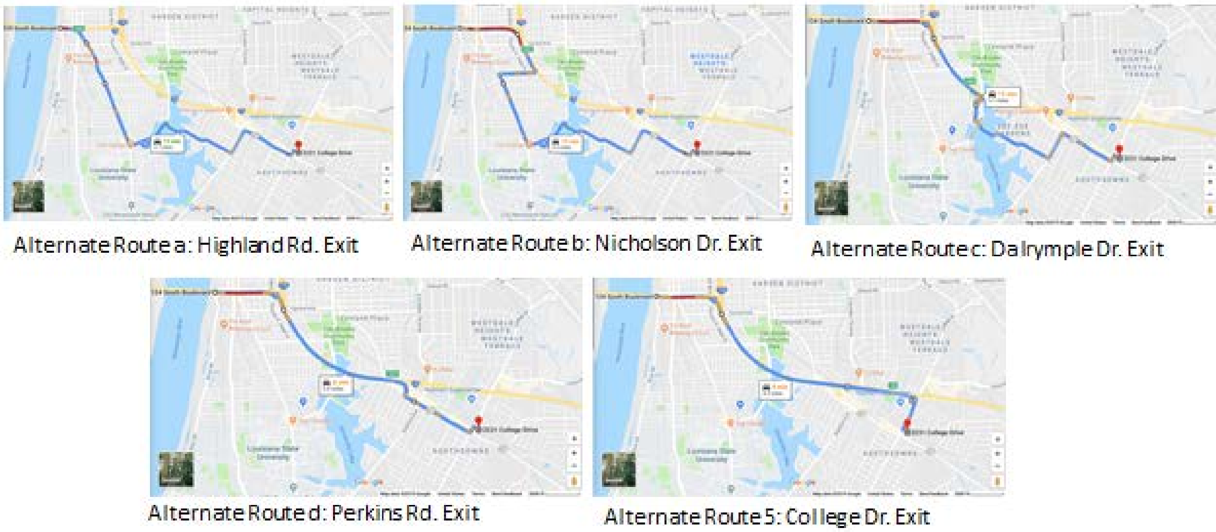

Figure 1. Maps showing alternative routes for the driving environment of the study.

4.2.2. VR Experimental Setting

In creating the experimental platform of this study, various types of cues were designed and used in order to convey the required information to put participants into the desired situations and allow them to engage the experiments. In fact, adding cues to the VR was used as a method for creating situation awareness (Figure 2) — the process of realizing the surroundings *(2).* The study employed a narrative-along-the-experiment to assist participants to follow all steps without the need to break their connection with the VR. Dynamic route guidance was presented to them where a certain driver is guided on to routes that will minimize travel time for the overall road network. For instance, about 200 ft. before reaching the exits, an audio message would be played to inform the participant about the chance (%) of reaching the destination on-time, if the forthcoming exit is taken. There were also other types of cues (e.g. clock, road signs (see Figure 3), traffic noise) to attract participants' attention to the spatial, temporal features of each of the scenarios. The experimental scenarios of this study are shown in Table 1.

Table 1. Experimental traffic condition scenarios used in the study.

Scenario #	Traffic Condition	Journey Type	Social Impact
1	Normal	Urgent-bound	No
2	Medium	Urgent-bound	No
3	Heavy	Urgent-bound	No
4	Medium	Urgent-bound	Yes
5	Heavy	Urgent-bound	Yes
6	Normal	Non-urgent-bound	No
7	Medium	Non-urgent-bound	No
8	Heavy	Non-urgent-bound	No
9	Medium	Non-urgent-bound	Yes
10	Heavy	Non-urgent-bound	Yes

4.2.3. Participants and Procedure

The study received approval from the IRB at Louisiana State University. Forty-one individuals (20 male and 21 females; age: 31.44 ± 7.97) volunteered to participate in the experiment, and no specific criteria was used for recruitment. Before the experiment started, participants signed the consent form and only after that the instructions about the experiment could be given to the participants. Prior to the experiment, participants were presented with a questionnaire asking the following items: (1) demographic characteristics (age, gender, race, education, employment status), (2) top concerns while they stuck in the traffic congestion. Their choices included hours of extra travel time, speed reduction, monetized value of delay, additional vehicle operating cost, (3) familiarity with the area, and (4) socio-economic status (having concerns about spending less money on your gas). After answering the questionnaires, participants were asked to sit on a stationary chair at a desk with a driving wheel which was placed in front of a 19 inch flat screen monitor where the driving simulation would run (Figure 2). Next, they were invited to practice for a few minutes to get acquainted with driving the simulator. After enough practicing with the driving simulator and becoming comfortable with its environment the research team would assign the participant to the scenarios. The scenarios were counterbalanced and played out in a random fashion to avoid behavioral biases due to order effect. The origin and the destination in all the scenarios were the same and each scenario took about two minutes to finish. In each scenario, different contextual factor(s) were presented and participants were required to choose their preferred route.

Figure 2. A volunteer using the driving simulator.

Figure 3. A screenshot from the VR experiment environment.

4.2.4. Data

The main variable of interest (dependent variable) in this study was the route choice while the contextual factors were considered as independent (manipulated) variables. Route choice is a categorical (nominal) variable with 5 levels—a, b, c, d, and e, and contextual factors shown in Table 1. In order to determine whether there is an association between the route choice and the contextual factors in this study, Chi-Square Independence Test is used. The study tries to refute the null hypothesis that two categorical variables are independent in this sample. This study used SAS 9.4 and JMP Pro 14 to perform the data analysis.

High fidelity route choice models are required to predict traffic levels with higher accuracy. Existing route choice models use revealed preference behavior to model route choice. The use of revealed choice data limits the accuracy of the prediction as it fails to capture subjective factors of drivers at individual level and allows prediction only at an aggregate level. Figure 4 shows the route choice predictions made by a basic aggregate level route choice model (blue line) compared with real data collected from the field (red line). The baseline model predicts the probability (P_b) of exiting a highway through a given exit using the following Equation 1.

$$P_b = \alpha_b T \qquad [1]$$

where:
the constant $\alpha_b = 0.601$ *(1);* and
T = reciprocal of the time needed to travel on the alternative route to a fixed destination after exiting the highway.

This model *(1)* is essentially based on the fact that drivers usually tend to choose the route with the least travel time (e.g., in a GPS). However, there is plenty of evidence that as commuters we take routes that do not minimize travel time *(30).*

More precisely, Figure 4 shows the probability of drivers exiting a freeway segment through one of the four available exits as predicted by the baseline aggregate route choice model in equation 1 (blue line); the red line in Figure 4 shows the ground truth. It can be seen from Figure 4 that the predictions made by the basic model deviate widely from the ground truth. Existing route choice models do not take into account dynamic contextual conditions such as the occurrence of an accident, the socio-cultural and economic background of drivers, other human behaviors, the dynamic personal risk level, etc. As a result, they can only make predictions at an aggregate level and for a fixed set of contextual factors. Therefore, for higher fidelity, it is highly desirable to use a methodology that captures significance of subjective or contextual factors in route choice.

Adding subjective or contextual requires availability of the data at individual or disaggregate level. Stated Choice Experiments (SCEs) are a scientific methodology to capture the effect of context sensitive factors in route choice. The current advancements in virtual reality technology can enhance stated choice experiments by allowing researchers to present them in a realistic manner that enhances the realism of the experiments and allows one to elicit information about route choice made by a driver. Interactive Virtual Environments (IVEs) provide a good platform to conduct SCE and elicit responses to route choice experiments as realistically as possible. The promise of IVE applications in collecting data includes, but is not limited to, providing a safe and user-friendly experimental platform, being inexpensive and highly portable, as well as generating context-aware and high-fidelity data.

4.2.5. High Fidelity Route Choice Models

The second goal of this study is to develop a novel approach for developing high-fidelity route choice models with increased predictive power by augmenting existing aggregate level baseline models with information on drivers' responses to contextual factors obtained from SCE carried out in an IVE through the use of knowledge distillation. Our approach uses the prior knowledge acquired by a teacher neural network pretrained on data about drivers' responses to contextual factors to augment a student neural network (a baseline model) in a guided fashion. We will demonstrate experimentally that the predictions of the augmented model are much closer to reality than that of the baseline.

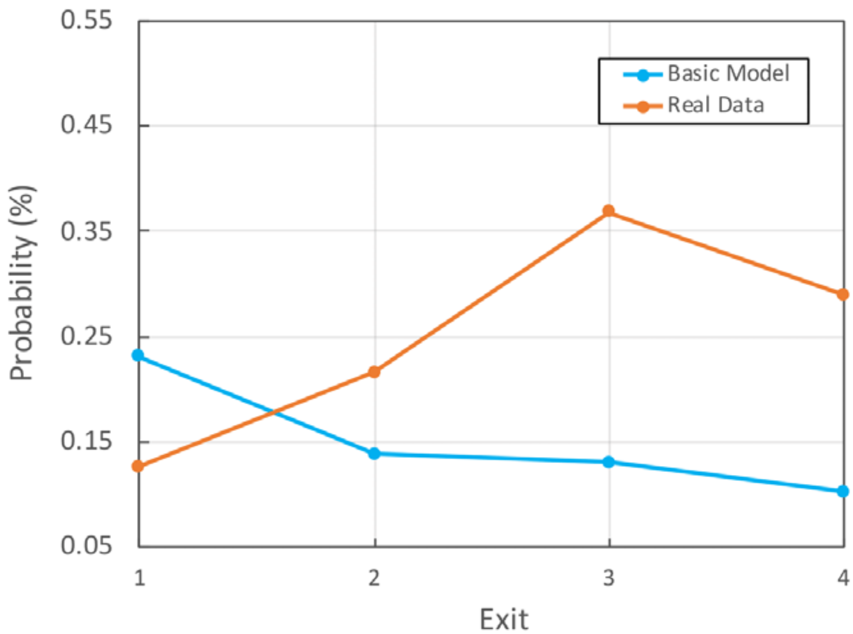

Figure 4. Comparison of predictions made by an aggregate route choice model (blue line) and ground truth (orange line).

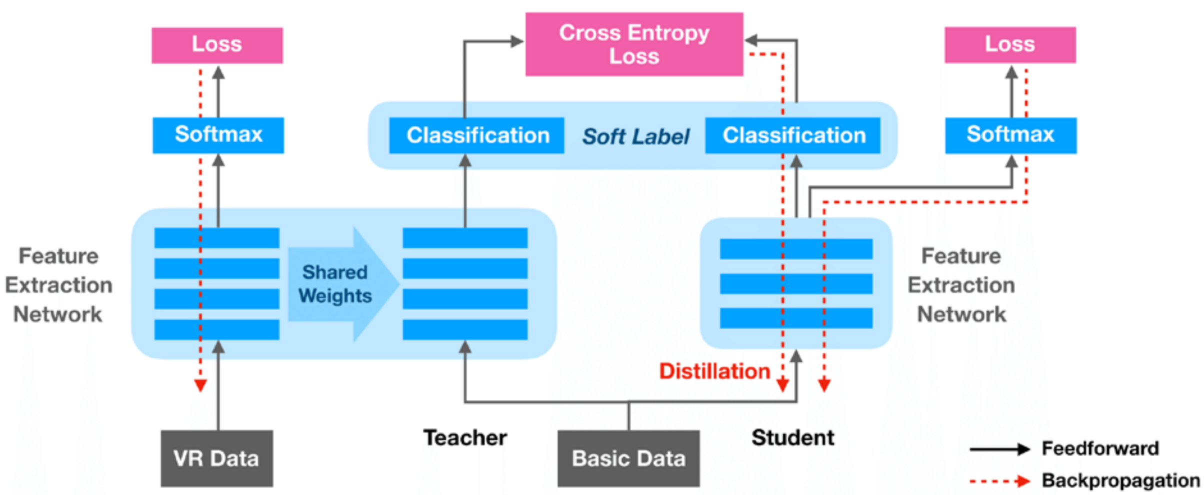

Figure 5. Architecture of the knowledge distillation framework.

The overall architecture of our framework is shown in Figure 5. As shown in the figure, for both the teacher and the student we use feature extraction networks. The teacher is first pre-trained on data acquired from SCE in IVE (called VR data). The basic data (see below) acquired from predictions by the baseline route choice model is partitioned into a training set and a test set. The training set is used for training the student as well as for distilling knowledge from the teacher to the student through knowledge distillation. The student is evaluated on the test set. During experimentation, we varied the number of the neurons in each layer activated by ReLU function as well as the number of Dropout and hidden layers. In each architecture, the output from the last layer of the feature extraction network is input into a 4-way softmax layer that transforms the logits to a probability distribution over the four exits. For student network, we implemented a small feature extraction network with two dense layers of 10 and 20 neurons, both layers activated by ReLU functions. No dropout layer was used in the student network. Instead, we added a Batch Normalization layer to its second dense layer, and a 4-way softmax layer on the top of the last layer. The inputs for the teacher and student networks have both 12 dimensions. In our experiments, we computed the softened logits output from the last layer of our feature extraction networks, and then the softened softmax output is obtained by applying the softmax layer on the softened logits. The original ground truth data concatenated with the softened softmax outputs from teacher network is used for training and testing the student network. The predictions from the student network concatenated with its softened softmax outputs are used to compute the distillation loss in each iteration. Then we use backpropagation for updating the parameters of the student network using gradient descent. Finally, the standalone trained student network is used for inference.

We considered a basic mathematical route choice model adapted from one of the common route choice models in literature *(1)* that serves as the baseline model. The route choice model is given in Equation 1 above. The highway segment chosen for our experiment corresponds to the route of I-10 in in Baton Rouge, Louisiana, United States, between Horace Wilkinson Bridge and the intersection of Perkins Rd and Staring Ln, that has four exits in the middle. The travel time measured on the alternative route after taking each exit provided by Google Maps 1 on September 20th, 2018, are 31:7 min, 18:9 min, 17:8 min, and 13:9 min. For the basic data, we uniformly randomly sampled 10,000 driving records based on the probability distribution predicted by the baseline route choice model. Each record was associated with its travel time corresponding to the alternative route for the exit taken by a driver. For each driving record, we assigned the value one to the Urgency variable if travel time is less than or equal to 13 (the threshold for urgency) in the scale of 1 to 60. Otherwise it was assigned to two. We did not consider the social impact factor in the normal traffic scenario (that is the impact on an individual driver on seeing a large number of drivers taking an exit) in our experiment due to the high cost in creating such a scenario.

From the SCE in IVE involving 41 volunteers, a total of 410 driving records collected. Since the data collected from SCE in IVE is limited, to better train the teacher network, we augmented it using a Gaussian mixture model (GMM). The data collected was categorical. We preprocessed and transformed it to the ordinal data before augmenting it using a GMM. After data augmentation, 10,000 synthetic driving records were generated. Each driving record was associated with its travel time corresponding to the alternative route for the exit taken by a driver. The 10,000 synthetic driving records together with their associated travel time is called the VR data. The VR data was divided by 80% for training and 20% for testing and the training set was used to train the teacher.

The contextual variables, which are present in the VR data, but do not occur in the basic data are set to zero. We used this augmented dataset for knowledge distillation. We divided this dataset into training (80%) and testing sets (20%).

During knowledge distillation, the teacher model, pretrained on the augmented VR data, provides the prior knowledge for our framework. During training the student model on basic data, our framework incorporated the pretrained teacher model for guiding the student. Specifically, it computes the cross-entropy loss in softened softmax function between the teacher and the student model in the backpropagation procedure for distilling the knowledge to student model. During inference, and we executed the student model as a standalone. It extracts features of each test data point from dense layers then predicts the route choice probability distribution.

We calculated the real probabilities of taking the exits from the data provided by LTRC. Given the traffic volumes captured at the four exits. The probability of taking an exit e is computed by,

$$P_e = \frac{V_e}{\sum_{i \in E} V_i} \quad [2]$$

where:
E is the set of exits; and
V_i is the traffic volume at exit i.

The real probabilities were used for evaluating the accuracy of the predictions by our framework.

5. ANALYSIS AND FINDINGS

The data is displayed in a mosaic plot (Figure 6), where each row represents a category of the route choice (a through e) and each column represents an experimental scenario, containing the mixture of the contextual factors. Figure 5 illustrates a mosaic plot in that it is clear that the two variables (a and e) are closely related to each other. Generally, when the variables are independent of one another, the proportions would be close and therefore the boxes would line up in a grid. In the case of this study, the structure of the frequency of the route choice across each experimental scenario significantly varies and the p-value associated with the chi-square test confirms this observation ($p<0.001$, $\alpha=0.05$). Thus, the route choice variable has an association with the contextual factors. Furthermore, the authors sought to investigate the impact of each of the contextual factor as well as the human-related variables, by developing a predictive statistical model *(29, 30)*. To that end, the logistic regression model was used since the dependent variable of the study was categorical. However, a specific type of logistic regression, namely, logistic regression with GEE (Generalized Estimating Equation *(29)*) was applied for this specific dataset for two main reasons: (1) the dependent variable of the study was dichotomous and a set of explanatory variables were available, and (2) the data were clustered, involved repeated measures. That is, 10 repetitions for each respondent were grouped as a cluster in a regression. Logistic regression with GEE was first introduced by Liang and Zeger in 1986 *(31)* and unlike the regular logistic regressions, this model allows for dependence within clusters, such as in longitudinal data.

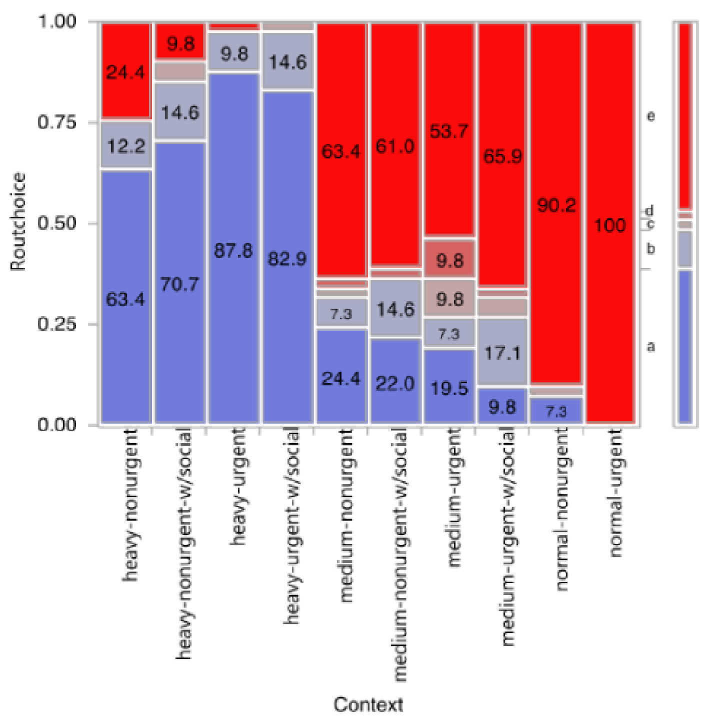

Figure 6. Frequencies of route choices under each experimental scenario.

Since the five original categories of the responses were unbalanced, the response variable was grouped into three new categories: (1) the first exit (a), (2) the interim exits (b, c, d), and (3) the last exit (e). The authors ran two separate regression analyses— nominal logistic regression and ordinal logistic regression. The ordinal logistic regression showed a better fit since it had a smaller QIC (Quasi-likelihood Information Criterion). The model equation of the ordinal logistic regression with GEE is as follows:

η_1= 4.68-6.33X_1-3.29 X_2-0.079 X_3+1.30 X_4; η_2= 5.91-6.33X_1-3.29 X_2-0.079 X_3+1.30 X_4

$$P(R = 1) = \frac{e^{\eta 1}}{1+e^{\eta 1}} \quad [3]$$

$$P(R= 2) = \frac{e^{\eta 2}}{1+e^{\eta 2}} - \frac{e^{\eta 1}}{1+e^{\eta 1}} \quad [4]$$

$$P(R= 3) = 1 - \frac{e^{\eta 2}}{1+e^{\eta 2}} \quad [5]$$

where:
$P(R=1)$, $P(R=2)$ and $P(R=3)$ are the estimated probabilities that the driver chooses to exit at the first exit, interim exits, and the last exit, respectively.

All independent variables except *Age* are categorical. The subscript number following the variable name means the specific category to the base category of that independent variable. The last category for each independent variable is set as the "base" for logits. Table 2 demonstrates the score statistics for the GEE analysis.

Table 2. Parameter estimates, standard error, p-value, CI95% for models.

Parameter	Coefficient	Estimate	Std. Error	$LCI_{95\%}$	$UCI_{95\%}$	*p*-value
Intercept 1	$\hat{\alpha}_1$	4.68	1.64	1.46	7.90	0.004
Intercept 2	$\hat{\alpha}_2$	5.91	1.62	2.73	9.09	0.0003
Traffic condition (X_1)	$\hat{\beta}_1$	-6.33	0.63	-7.5670	-5.09	<0.0001
Traffic condition (X_2)	$\hat{\beta}_2$	-3.29	0.32	-3.90	-2.67	<0.0001
Age (X_3)	$\hat{\beta}_3$	-0.08	0.03	-0.13	-0.02	0.007
Race (X_4)	$\hat{\beta}_4$	1.3006	0.6264	0.0728	2.5285	0.0379

The score statistics for the type 3 GEE analysis indicate that the impact of "traffic condition," "age," (age distribution quantile plot provided in Figure 7) and "race" (White, Hispanic or Latino, Middle Eastern, Asian; ethnicity distribution provided in Figure 8) are statistically significant (*p*-value<0.05), meaning that these three variables can significantly affect individuals' route choice behavior. More specifically, parameter estimates of the aforementioned variables show that when other variables are fixed, (1) participants (driver) tend to drive longer distances before exiting I-10 if the traffic condition is better, (2) older people tend to drive longer distances before exiting I-10, and (3) a driver of Race=2 (Middle Eastern) tends to drive fewer distances before exiting I-10 compared to a driver of Race=3 (Asian) *(32,33).*

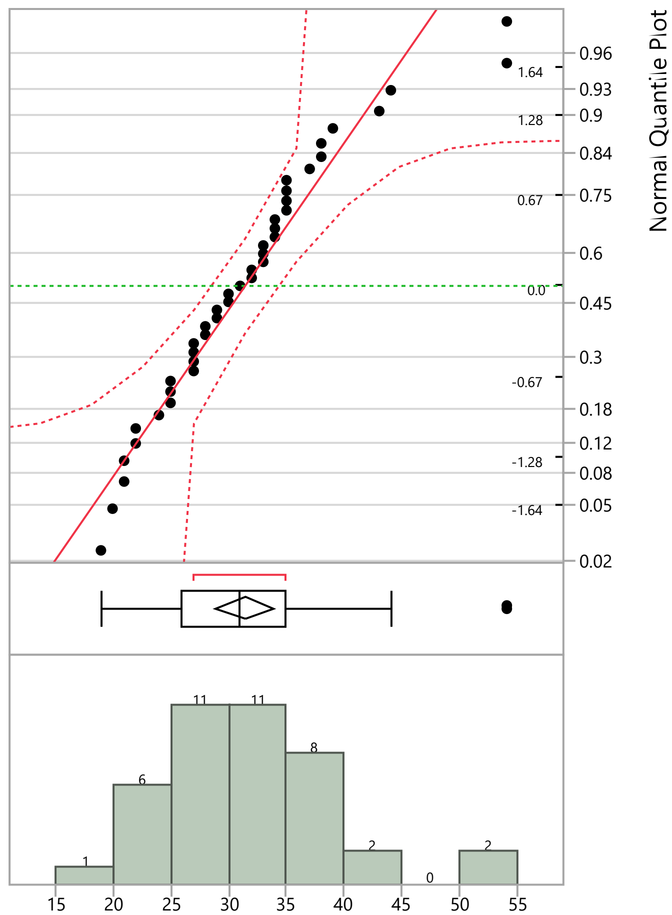

Figure 7. Age distribution of participants in the study.

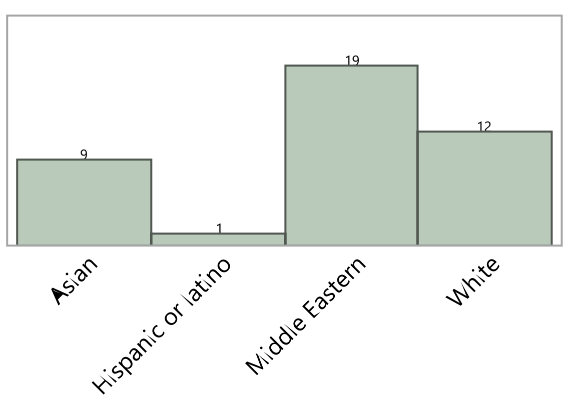

Figure 8. Ethnicity distribution of participants in the study.

5.1. Machine Learning Model

We developed a novel approach *(34)* for building high-fidelity route choice models with increased predictive power by augmenting existing aggregate level baseline models with information on drivers' responses to contextual factors obtained from stated choice experiment (SCE) carried out in an IVE through the use of knowledge distillation. Our approach uses the prior knowledge acquired by a teacher neural network pre-trained on data about drivers' responses to contextual factors to augment a student neural network (a baseline model) in a guided fashion. After knowledge distillation from the teacher network to the student network, the prediction accuracy of the student network on the test set of the basic data abruptly improves: it achieves a classification accuracy of 95.2% on the basic data. We demonstrated experimentally that the predictions of the augmented model are much closer to reality than that of the baseline (see Fig. 9). It can be seen from Figure 9 that the prediction accuracy of our framework is better than the baseline model: our results are closer to the real data and have similar trend except at Exit 1. The prediction at Exit 1 was heavily dominated by a large number of discrepancies in drivers' actions as seen from the VR data. However, overall, our framework shows better fidelity than the baseline route choice model. Thus, using knowledge distillation, we have augmented a baseline model with contextual information acquired from SCE to obtain a model with higher fidelity.

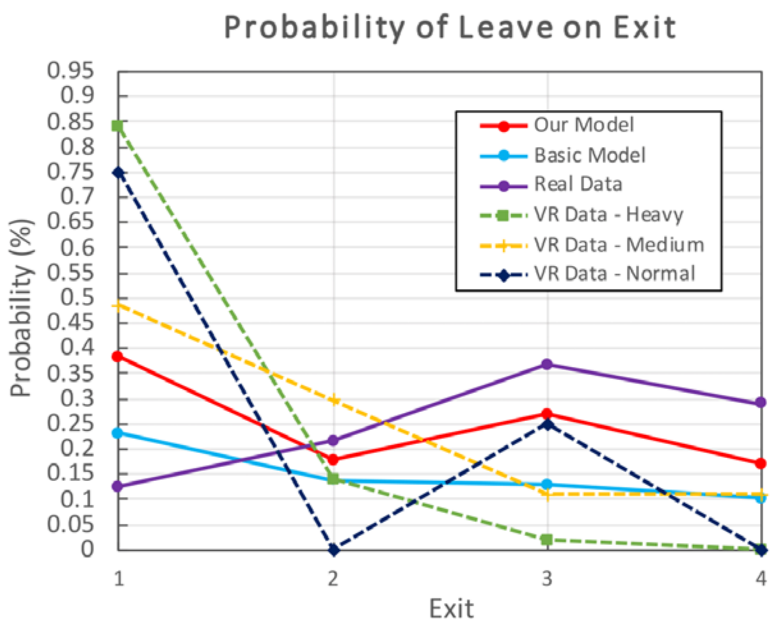

Figure 9. Comparison of route choice probability obtained from augmented model (red line) with (a) baseline model (blue line), (b) real data (purple line).

5.2 Limitations of the Study and Possible Future Work

Even though the results of the study are promising, IVE-based SCE's have their inherent limitations. Due to limitations of IVE technologies, it is difficult to continuously collect data on interaction of humans with transportation systems. Our current work is exploring the use of Shannon-Nyquist sampling techniques to determine the optimal rate of sampling for effective inference. Due to the expensive nature of the IVE SCEs, we could only explore a limited set of scenarios. While it showed the potential of the framework, there is scope for computational bias and overfitting. We have attempted to reduce the size of the model using knowledge distillation to counter overfitting. However, more scenarios need to be considered. In addition, bias-variance-cost trade-offs need to be investigated. Also, due to the time and cost limitations, only a small sample size of participants (41) was considered. To this end, low shot and unsupervised learning techniques *(35)* need to be considered to learn from a small dataset. In addition, in future, we will consider more variation in the set of participants, e.g., people who do not speak any English. Also, in this study, we considered a simplistic, but often used (especially in GPS-based systems), baseline model *(1)*. In the future, we will richer baseline models. In addition, in the future, we will consider a richer IVE framework depicting critical events such as storms, etc. We will also consider modeling evacuations in hurricane/storm scenarios. We have also observed that there is not much research that have tried to apply causal analysis methods to explain route choice behavior. We have applied causal analysis techniques to identify root causes that influence route choice *(36).* We believe that this will subsequently allow us to enhance Route Choice models with more appropriate context factor, such as, driving experience, that will better forecast traffic levels on transportation networks and also to better comprehend drivers' response to route guidance and dynamic message signs.

6. CONCLUSIONS

The promise of VR application in collecting data includes, but is not limited to, providing a safe and user-friendly experimental platform, being inexpensive and highly portable, while generating context-aware and high-fidelity data. While most of the conventional route choice models use aggregated data, this study provides the opportunity to incorporate any contextual factor in the experiment, and include human-related factors, as well. The main contextual factors experimented in this study were traffic condition, journey type, and social impact. Furthermore, participants' demographic information (i.e. age, gender, race, education, employment status), familiarity with the area, top concern while stuck in traffic and their financial concerns were gathered. Results of the data analysis revealed that "traffic condition," "age," and "race" were statistically significant in this study. However, since the sample of this study may not be fully representative, the results might not have a strong external validity. Therefore, the main contribution of this study will be introducing a new data collection method for traffic-related studies.

In this study, we also proposed a novel approach for developing high-fidelity route choice models with increased predictive power by augmenting existing aggregate level models with contextual information obtained from SCE carried out in an IVE through the use of knowledge distillation. To this end, we presented a general end-to-end knowledge distillation framework that uses a multilayer perceptron as a feature extraction network to provide a feature learning architecture for teacher and student networks and then transfers knowledge from the former to the latter by optimizing distillation loss. Experimental results have shown that that the predictions of the augmented models produced by our approach are much closer to reality than that of the baseline.

The study demonstrated that route choice models based on econometric theories do not accurately reflect ground truth. These models are based on aggregate behavior and do not take into account context-sensitive factors that influence decision-making of individual drivers. For high-fidelity route choice models, one needs to combine existing route choice models with information about contextual factors gleaned from SCEs.

REFERENCES

1. Ben-Akiva, M., and M. Bierlaire. “Discrete choice models with applications to departure time and route choice.” *Handbook of transportation science*, Springer, 2003, 7–37.

2. Endsley, M.R. . “Situation awareness global assessment technique (SAGAT).” *Proceedings of the Institute of Electrical and Electronics Engineers 1988 National Aerospace and Electronics Conference*, 789–795.

3. Schrank, D., Lomax, T., and S. Turner. “Urban mobility report texas transportation institute.” *Texas: Texas Transportation Institute, 2009*.

4. Gudishala, R., Wilmot, C., and M.R.V. Peddisetty. *Diverted Traffic Measurement*. *Unpublished*.

5. Shojaat, S., Geistefeldt, J., Parr, S. A., Wilmot, C.G., and B. Wolshon. “Sustained Flow Index: Stochastic Measure of Freeway Performance.” *Transportation Research Record: Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, 2016, (2554), 158–165.

6. Prato, C.G. “Route choice modeling: past, present and future research directions.” *Journal of choice modelling*, Leeds: Institute for Transport Studies, University of Leeds, 2009, 2(1), 65–100.

7. Ben-Akiva, M., S.R. Lerman, and S.R. Lerman. Discrete choice analysis: theory and application to travel demand. MIT press, 1985, vol. 9.

8. Machado-León, J. L., de Oña, J., de Oña, R., Eboli, L., and G. Mazzulla. “Socio-economic and driving experience factors affecting drivers’ perceptions of traffic crash risk.” *Transportation research part F: traffic psychology and behaviour*, Elsevier, 2016, 37, 41–51.

9. M. Pursula and A. Talvitie, “Urban route choice modelling with multinomial logit models,” LIIKENNETEKNIIKKA, TIEDOTE, 1993, no. 28.

10. A.J. Khattak, J.L. Schofer, and F.S. Koppelman, “Commuters’ enroute diversion and return decisions: analysis and implications for advanced traveler information systems, *Transportation Research Part A: Policy and Practice*, 1993, vol. 27, no. 2, pp. 101–111.

11. P.H. Bovy and E. Stern. Route Choice: Wayfinding in Transport Networks: Wayfinding in Transport Networks. *Springer Science & Business Media*, 2012, vol. 9.

12. S.T. Doherty and E.J. Miller, “A computerized household activity scheduling survey,” *Transportation*, 2000, vol. 27, no. 1, pp. 75–97.

13. Song, X., Kanasugi, H. and R. Shibasaki. Deeptransport: Prediction and simulation of human mobility and transportation mode at a citywide level. In *Ijcai* 2016, (Vol. 16, pp. 2618–2624).

14. Taheri, S. M., Matsushita, K., and M. Sasaki. “Development of a Driving Simulator with Analyzing Driver’s Characteristics Based on a Virtual Reality Head Mounted Display.” *Journal of Transportation Technologies*, Scientific Research Publishing, 2017, 7(03), 351

15. Saeidi, S., Zhu, Y., Chokwitthaya, C., and M. Sun. “Spatial-Temporal Event-Driven Model for Occupant Behavior Studies Using Immersive Virtual Environments : A Case Study.” *Journal of Computing in Civil Engineering*, 2017.

16. Jeihani, M., Narooienezhad, S., and K. Bakhsh. “Integration of a driving simulator and a traf fi c simulator case study : Exploring drivers ’ behavior in response to variable message signs.” *IATSS Research*, International Association of Traffic and Safety Sciences, 2017, 41(4), 164–171.

17. Jin, M., and S.-H. Lam. “A virtual-reality based integrated driving-traffic simulation system to study the impacts of intelligent transportation system (ITS).” *Cyberworlds, 2003. Proceedings. 2003 International Conference on*, IEEE, 158–165.

18. Heydarian, A, Carneiro, J. P., Gerber, D., and B. Becerik-Gerber. “Immersive virtual environments, understanding the impact of design features and occupant choice upon lighting for building performance,” *Build. Environ.*, 2015vol. 89, pp. 217–228, Jul.

19. Saeidi, S, Rizzuto, T., Zhu, Y., and R. Kooima. “Measuring the Effectiveness of Immersive Virtual Environment for the Modeling and Prediction of Occupant Behavior,” in the *1st International Symposium on Sustainable Human and Building Ecosystems*, 2015.

20. Niu, S., Pan, W., and Y. Zhao. “A virtual reality integrated design approach to improving occupancy information integrity for closing the building energy performance gap,” *Sustain. Cities Soc.*, vol. 27, pp. 275–286, 2016.

21. Hülsmann, F., Mattar, N., Fröhlich, J., and I. Wachsmuth. “Simulating wind and warmth in virtual reality: conception, realization and evaluation for a CAVE environment.” *JVRB - Journal of Virtual Reality and Broadcasting*, 2014, 11(10)

22. Zhu, Y., Saeidi, S., Rizzuto, T., Roetzel, A., and R. Kooima. “Potential and challenges of immersive virtual environments for occupant energy behavior modeling and validation: A literature review,” *J. Build. Eng.*, vol. 19, 2018.

23. Ben-Elia, E. and Y. Shiftan. Which road do i take? A learning-based model of route-choice behavior with realtime information. *Transportation Research Part A: Policy and Practice*, 2010.

24. Basu, S., Ganguly, S., Mukhopadhyay, S., DiBiano, R., Karki, M., and R.R. Nemani. “Deepsat: a learning framework for satellite imagery". In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Bellevue, WA, USA, November 3-6, 2015, 2015, pp. 37:1–37:10.

25. Basu, S., Karki M., Mukhopadhyay, S., Ganguly, S., Nemani, R. R., DiBiano, R., and S. Gayaka. A theoretical analysis of Deep Neural Networks for texture classification. *IEEE International Joint Conference on Neural Networks (IJCNN)* 2016: 992-999.

26. Basu, S., Ganguly, Nemani, R. R., Mukhopadhyay, S., Zhang, G., Milesi, C., Michaelis, A. R., Votava, P., Dubayah, R., Duncanson, L., Cook, B.D., Yu, Y., Saatchi, S., DiBiano, R., Karki, M., Boyda, E., Kumar, U., and L. Shuang. A Semiautomated Probabilistic Framework for Tree-Cover Delineation From 1-m NAIP Imagery Using a High-Performance Computing Architecture. *IEEE Trans. Geoscience and Remote Sensing*

53(10): 5690-5708 (2015)

27. Nguyen, H.,Kieu, L., Wen, T, and C. Cai. “Deep learning methods in transportation domain: a review”. *IET Intelligent Transport Systems*, vol. 12, no. 9, pp. 998–1004, 2018.

28. Krizhevsky, A., Sutskever, I., and G.E. Hinton. “Imagenet classification with deep convolutional neural networks”. In *Advances in neural information processing systems*, 2012, pp. 1097–1105.

29. Collier, E., DiBiano, R., and Mukhopadhyay, S. CactusNets: Layer Applicability as a Metric for Transfer Learning. *IJCNN* 2018: 1-8.

30. Ben-Akiva, M. E., Ramming, M. S., and S. Bekhor. Route choice models. In H*uman behaviour and traffic networks*. Springer, 2004.

31. Liang, K.-Y., and S.L. Zeger. “Longitudinal data analysis using generalized linear models.” *Biometrika*, Oxford University Press, 1986, 73(1), 13–22.

32. Saeidi, S., Zhu, Y., Mukhopadhyay, S. Gudishala, R., and Z. Xu. Application of Virtual Reality to investigate Driver's Route Choice in an Interstate Highway. In Proceedings of *2019 ASCE International Conference on Computing in Civil Engineering*.

33. Saeidi, S., Zhu, Y., Mukhopadhyay, S., Gudishala, R., and Z. Xu. Application of Virtual Reality to investigate Driver's Route Choice in an Interstate Highway. Abstract presented at *Tran-SET Conference* 2019.

34. Liu, Q., Mukhopadhyay, S., Zhu, Y., Gudishala, R. Saeidi, S. and A. Nabijiang. Improving Route Choice Models by Incorporating Contextual Factors via Knowledge Distillation, In *Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN)*, 2019.

35. Liu, Q., Mukhopadhyay, S. Unsupervised Learning using Pretrained CNN and Associative Memory Bank. In *Proceedings of IEEE International Joint Conference on Neural Networks* IJCNN 2018: 1-8.

36. Nabijiang, A., Mukhopadhyay, S., Zhu, Y., Gudishala, R., Saeidi, S., and Q. Liu. Why do you take that route? In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*. 2019.